

Eur. J. Clin. Chem. Clin. Biochem.  
Vol. 31, 1993, pp. 139–146

© 1993 Walter de Gruyter & Co.  
Berlin · New York

## Interlaboratory Surveys of the Determination of Tumour Markers Scatter and Repeatability of the Results

By R. Kruse, W.-J. Geilenkeuser and G. Röhle

*Institut für Klinische Biochemie der Universität Bonn*

(Received July 9/November 30, 1992)

**Summary:** Data collected between 1983 and 1991 in interlaboratory surveys of the determination of tumour markers are used to show the magnitude of the scatter of results from different laboratories for the analysis of a single quantity in a given matrix. These data also show that the varying specificity of different reagent combinations appears to make a considerable contribution to this scatter, and that the used reagent combinations were not of uniform quality over a relatively extended period. The results for the following tumour markers were studied:  $\alpha$ -fetoprotein (AFP), carcinoembryonic antigen (CEA), human chorionic gonadotropin (hCG), human chorionic gonadotropin +  $\beta$ -subunit (hCG +  $\beta$ -hCG), tissue polypeptide antigen (TPA), carbohydrate antigen 19-9 (CA 19-9), cancer antigen 15-3 (CA 15-3), cancer antigen 125 (CA 125), prostatic acid phosphatase (PAP) and prostate-specific antigen (PSA).

### Introduction

During the last decade, in the field of tumour diagnosis, many more quantitative determinations for tumour markers have become available, and the number of these determinations requested and performed has also markedly increased. The determination methods are based exclusively on immunochemical principles, whose complex procedures are more subject to systematic and random errors, compared with many procedures of more traditional clinical chemical analysis.

Soon after the introduction of the determination of tumour markers, the necessity arose to test the reliability of the analytical results in collaborative interlaboratory surveys, as far as this was possible under the preconditions laid down for such tests (1). Since 1981, regular interlaboratory surveys of tumour markers have been carried out in collaboration with the German Society for Clinical Chemistry. The first interlaboratory surveys were conducted on three tumour markers:  $\alpha$ -fetoprotein, carcinoembryonic antigen and human chorionic gonadotropin (2). By 1991 the

number of tumour markers under survey had increased to 10 and, by the end of 1991, the number of participating laboratories to 430.

More thorough investigations were promoted by the survey results of 1991, which revealed an especially wide scatter for certain analytical quantities.

The way in which the concept of the surveys is put into practice provides no basis for a detailed criticism of methods. Rather, these investigations had the following two-fold aim: on the one hand they sought to establish the magnitude of the scatter of results from different laboratories for individual tumour markers during an extended observation period. On the other hand, it was of interest to determine whether any characteristic differences existed between results obtained with different reagent combinations (kits). The results from the most frequently determined tumour markers, carcinoembryonic antigen and  $\alpha$ -fetoprotein, provide the most comprehensive data for the discussion of this problem. It therefore seemed reasonable to base a more detailed study on these two analytical

quantities. In this connection, the results of surveys, in which samples of the same material were repeatedly analysed at different time intervals, were especially informative.

## Materials and Methods

### Collaborative interlaboratory surveys

Two pilot surveys were conducted in 1981 and 1982. Since 1983, three interlaboratory surveys have been performed each year. The results of the surveys from 1983 to 1991 inclusive are the subject of the present communication. Two different samples for the quantitative determination of tumour markers were used in each survey. Aliquots of these samples were sent to each participating laboratory. The results were collected and evaluated by the survey management.

### Analytical quantities of the interlaboratory surveys

The surveys covered the following quantitative analyses for tumour markers:  $\alpha$ -fetoprotein (AFP), carcinoembryonic antigen (CEA), human chorionic gonadotropin (hCG), human chorionic gonadotropin +  $\beta$ -subunit (hCG +  $\beta$ -hCG), tissue polypeptide antigen (TPA), carbohydrate antigen 19-9 (CA 19-9), cancer antigen 15-3 (CA 15-3), cancer antigen 125 (CA 125), prostatic acid phosphatase, prostate-specific antigen.

### Sample material of the interlaboratory surveys

Sample material was prepared in our own laboratory. To increase the concentration of tumour markers, commercial preparations and/or pooled patient sera were added to pooled human serum. After homogenization and filtration (filter pore size 0.2  $\mu$ m), the resulting serum was divided into aliquots of 2.5 ml, which were placed in glass vessels, lyophilized and sealed.

Some examples of the sources of added tumour markers are: carcinoembryonic antigen (patient serum and commercial preparations from Calbiochem, PAN-Systems, and Dako Diagnostic);  $\alpha$ -fetoprotein (umbilical serum); human chorionic gonadotropin (Boehringer Mannheim); tissue polypeptide antigen (Sangtec), carbohydrate antigen 19-9 (patient serum), cancer antigen 15-3 (Centocor and patient serum), cancer antigen 125 (Centocor and patient serum), prostatic acid phosphatase (patient serum and commercial preparations from Scripps Laboratories), prostate-specific antigen (patient serum and commercial preparations from Scripps Laboratories).

### Evaluation of the interlaboratory surveys

The results of the evaluation of each analytical quantity were presented as follows.

1. A Youden diagram with identification of all single values, and separate labelling of the values of each participant.
2. The median and scatter (given as the 16th and 84th percentiles) for the total collective of all results for each analytical quantity.
3. The median and the 16th and 84th percentiles for each subcollective of users of the same method or commercial kit.
4. Graphical presentation of the data from 2. and 3. (see fig. 2).

Variability of the results for an analytical quantity in a survey sample

Variability is defined here as half of the difference between the 84th and 16th percentiles (P), relative to the median (as a percentage):

$$(84 P - 16 P) \times 100 / (2 \times \text{median}) = \text{variability.}$$

*Note:* In the unusual event that the results are normally distributed, the median is equal to the average value, and the variability is equal to the variation coefficient.

### Calculation of the average, concentration-dependent variability

Average scatters of the results of interlaboratory surveys can be displayed as concentration-dependent scatter profiles (3) (see fig. 3). For each survey collective, these profiles were constructed by plotting the relative differences between the median and the respective 16th and 84th percentiles against the median. Two hyperbolic regression curves (the scatter profiles) were then calculated, one for the upper and the other for the lower accumulation of points.

The variability data displayed in table 1 are based on survey data for the years 1986 to 1991; during this time 19 different specimens were used in the surveys. For quantities introduced later into the survey, fewer different specimens were used, i.e. 18 each for cancer antigen 15-3 and carbohydrate antigen 19-9, and 15 each for cancer antigen 125, prostatic acid phosphatase and prostate-specific antigen. The scatter profiles, from which the data are taken, were of two types. Thus, they were either calculated from the total collective of survey results for each single analytical quantity, irrespective of the analytical method, or they were calculated from subcollectives of results obtained with the same kit.

### Selected examples

A full and detailed treatment of all the results of the relevant interlaboratory surveys would exceed the reasonable limits of this publication. Examples have been chosen that are representative of the present state of development of tumour diagnosis, and for which the number of results within each collective is sufficiently large to ensure that the resulting data will be meaningful.

### Definitions of terms and codes

- |                   |   |
|-------------------|---|
| Kit:              | a product from a manufacturer or supplier of diagnostic reagents, offered commercially as an analytical method, and consisting of the instructions for use, the reagents and, where applicable, the analytical apparatus. This is the usual definition of a kit.    |
| Subcollective:    | all the analytical results for a given analytical quantity in a given matrix, determined with same kit.   |
| Total collective: | all the results for a given analytical quantity in a given matrix, irrespective of the methods used.  |
| Kit codes:        | Kits were assigned a three digit code. The last two digits identified the manufacturer, while the first digit indicated the analytical principle of measurement:<br>1: measurement of radioactivity<br>2: measurement of enzyme activity<br>3: measurement of light |

Kit 2 28 is therefore an enzyme-immunoassay from manufacturer 28.

## Accuracy

It is not possible to test the accuracy of tumour marker analyses, because no analytical methods (e.g. reference methods) are known for the reliable determination of the true tumour marker concentrations. Studies must therefore be based on the detection and evaluation of differences between individual analytical results or between groups of analytical results for analyses performed on a given material.

## Results and Discussion

### Number of participants

Figure 1 shows the increase in the total number of participating laboratories, as well as number of laboratories analysing each of eleven analytical quantities.

The total number of participants increased from 120 in 1983 to 430 in 1991. Whereas the number of participants for human chorionic gonadotropin (total), prostatic acid phosphatase and tissue polypeptide antigen almost stagnated, for the seven other analytes there was an increase in the number of participating laboratories, which was roughly parallel to the increase in total participants. There was a decrease in the number of laboratories that determined human chorionic gonadotropin and calibrated their assay

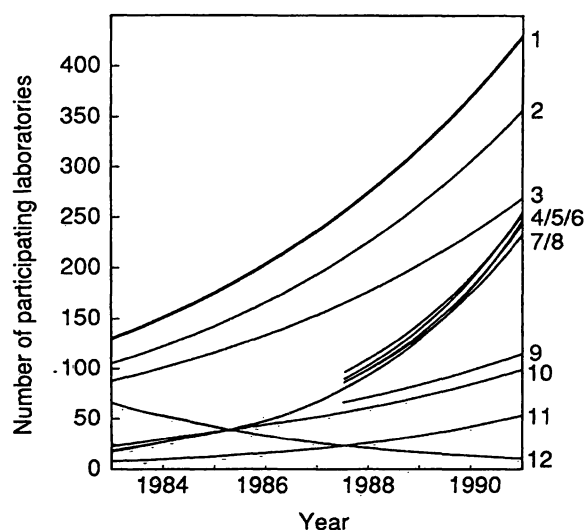


Fig. 1. Changes in the numbers of laboratories participating in interlaboratory surveys from 1983 to 1991.

- 1 = Total number of participants,
- 2 = carcinoembryonic antigen,
- 3 =  $\alpha$ -fetoprotein,
- 4 = carbohydrate antigen 19-9,
- 5 = prostate-specific antigen,
- 6 = cancer antigen 125,
- 7 = cancer antigen 15-3,
- 8 = human chorionic gonadotropin +  $\beta$ -subunit (1. IRP),
- 9 = prostatic acid phosphatase,
- 10 = tissue polypeptide antigen,
- 11 = human chorionic gonadotropin,
- 12 = human chorionic gonadotropin +  $\beta$ -subunit (2. IS)

with the second international standard for bioassays. Since participation is voluntary, the commitment of most of the analysts involved is probably above average. These numbers therefore provide only a qualified assessment of the overall situation. However, the continual increase in participants certainly indicates an increasing interest in external quality control, as well a real increase in the number of laboratories involved in the determination of tumour markers.

### Scatter of interlaboratory survey results

In evaluating the scatter of interlaboratory survey results, account must be taken of both internal laboratory scatter, and scatter between different laboratories. Internal laboratory precision depends primarily on the size of unavoidable, random errors. The scatter of results between different laboratories, as encountered in interlaboratory surveys, is additionally due to laboratory-specific differences in the implementation of methods, and especially to the marked dependency of the analytical results on the kit used. The variations caused by these two influences have the character of systematic errors. There is, of course, only one correct value for each analysis, but in the present state of development of tumour marker analysis, it is not possible to determine the systematic error for the determination of any tumour marker. Since reliable accuracy criteria (reference method values or assigned values) are lacking, it is not possible to determine which results are correct.

Principle	Kit manufacturer	N	CEA [ $\mu\text{g/l}$ ]					
			Min	16th percentiles	50th	84th	Max	
All		360	8.0	13.1	17.2	22.7	32.9	— + —
1	04	29	14.0	17.1	18.5	19.5	21.4	+ —
1	20	26	12.7	13.9	15.3	16.5	23.1	+ —
1	30	22	11.7	19.7	22.8	24.4	30	— + —
1	32	10	13.0		29.5		32.9	— —
1	72	8	22.0		25.0		25.6	— —
1	76	14	10.9	11.3	12.6	13.9	16.4	+ —
1	77	10	11.3		16.5		18.3	— —
2	04	52	8.0	12.2	13.4	15.0	21.1	+ —
2	28	63	13.0	16.3	17.5	18.5	21.2	— + —
2	74	23	12.3	18.9	22.7	25.2	28.1	— + —
3	04	25	12.0	12.4	13.0	14.2	14.9	+ —
3	08	18	15.2	18.0	18.7	20.2	22.4	— + —
3	32	11	11.2		23.8		26.8	— —

Fig. 2. Results ( $\mu\text{g/l}$ ) of the carcinoembryonic antigen determination in the interlaboratory survey of November 1991, subdivided according to methods and manufacturers, with graphical presentation of the medians (|) and the ranges between the 16th and 84th percentiles (—).

N = number of participants; Min = lowest analytical value; Max = highest analytical value; CEA = carcinoembryonic antigen

The graphical presentation of the medians and percentiles of an interlaboratory survey of the most frequently investigated analytical quantity, carcinoembryonic antigen (fig. 2), clearly shows that the total scatter of the results is crucially and greatly influenced by the position, i.e. the medians, of the individual subcollectives.

The medians of the subcollectives sometimes show considerable differences amongst themselves. Such differences are largely responsible for the large scatter of the values of the total collective (fig. 2, line "All"). The scatter of each subcollective (fig. 2, subsequent lines) is usually relatively small.

Table 1 lists the scatter of results for ten tumour markers, showing the values for total collectives, as well as those for subcollectives of the most frequently used kits. The following brief comments may be made on the origin and significance of these values.

The average variability, as derived from each concentration-dependent scatter profile, is the measure of scatter. As an example, figure 3 shows the scatter profile for the total collective of results for  $\alpha$ -fetoprotein; the regression curves show an increase of scatter with decreasing concentration precisely in the clinically relevant decision range. To define this situation briefly and numerically, the variability at the upper and lower limit of a selected concentration range was taken from the graph and listed in table 1. For the total collective of the  $\alpha$ -fetoprotein determination, the variability was 24% at 6 kU/l, and 18% at 30 kU/l. Corresponding scatter profiles were calculated for the other tumour markers, as well as for a series of relatively large subcollectives. Variability data were again taken from these profiles and listed in table 1.

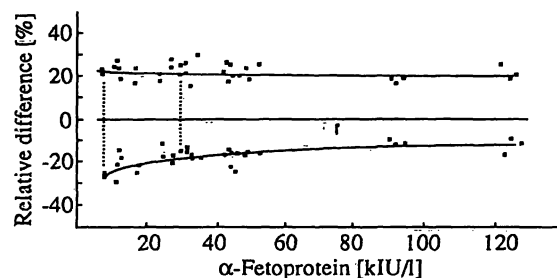


Fig. 3. Concentration-dependent scatter profiles (d(%)) (see Materials and Methods) for the analytical quantity  $\alpha$ -fetoprotein. The variabilities at two concentrations (6 and 30 kU/l) given in table 1 are labelled. Results for the period 1986–1991.

These variability data show

- the extent of the scatter that can be expected in the results from different laboratories, with the present state of development of tumour marker analysis,
- that the scatter, both of the total collective and of subcollectives, in most cases increased with decreasing analyte concentration,
- that the scatter of each subcollective, in analogy with the example in figure 2, was usually significantly less than in the corresponding total collective.

#### Kit-dependency of the results

The finding that the scatter displayed by a total collective was usually larger than that of each of its subcollectives suggested a possible causal relationship between the kit used and the values obtained. This was therefore subjected to a more detailed investigation, making use of available examples.

Tab. 1. Variability (see Materials and Methods) of the results of interlaboratory surveys for 10 tumour markers each of them analysed once or twice in up to 19 different specimens. Variabilities (for the total collective and for up to four subcollectives of laboratories using the same kit) were taken from the appropriate concentration-dependent scatter profiles (e.g. fig. 3) at two defined concentrations.

Analyte in serum	Concentration range of the samples used	Observed concentration range	Variability								
			All results	Results from users of defined kits							
				%	%	(Kit)	%	(Kit)	%	(Kit)	%
AFP	6– 134 kU/l	6– 30 kU/l	24–18	14–12	(1 76)	21–14	(2 04)	19– 8	(2 28)		
CEA	5– 52 µg/l	5– 30 µg/l	22–22	10–10	(1 04)	15–11	(2 04)	15– 9	(2 28)		
hCG	3– 180 U/l	5– 30 U/l	30–11								
hCG + β-hCG	5– 230 U/l	5– 30 U/l	30–11	24– 7	(1 76)	11–11	(2 04)				
TPA	94–4500 U/l	95–400 U/l	14–14	14–14	(1 32)						
CA 125	8– 41 kU/l	10– 40 kU/l	31–18	25–12	(1 04)	24–11	(1 30)	24–11	(1 32)	65–22	(2 28)
CA 15-3	11– 40 kU/l	13– 35 kU/l	16–16	15–15	(1 30)	14–14	(1 32)	14–14	(2 28)		
CA 19-9	8– 540 kU/l	16– 50 kU/l	35–26	21–14	(1 04)	18–13	(1 30)	25–21	(1 32)	11– 9	(2 28)
PSA	1– 12 µg/l	1– 10 µg/l	45–36								
PAP	1– 21 µg/l	2.5– 20 µg/l	33–26								

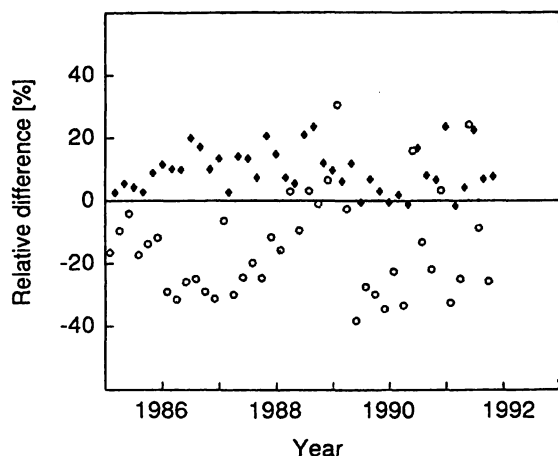


Fig. 4. Relative differences between the medians of the results for carcinoembryonic antigen obtained with kit 1 04 (◆) or kit 1 76 (○) and the corresponding medians of the total collective. Interlaboratory surveys (6 samples per year) conducted from 1986 to 1991.

Figure 4 shows, for the analytical quantity carcinoembryonic antigen, the relative deviations of the medians of the subcollectives kit 1 04 and kit 1 76 from the median of the total collective during the period of the investigation (two medians are not shown, because their values lay outside the presentation area). In most cases, the medians of kit 1 04 and kit 1 76 differed by between 30 and 40%. Relatively large differences between the medians of individual subcollectives are also shown in figures 5a and 5b for the determination of  $\alpha$ -fetoprotein.

Several reasons may be considered for the sometimes pronounced dependency of analytical results on the kit used. In addition to the antigen heterogeneity of native tumour markers (4), the multiplicity of commercial antibodies, some of them raised against different epitopes of the antigen, inevitably leads to differences in the composition of the kits. These differences probably make a considerable contribution to the observed kit-dependency of the results. The comparability of analytical results from different laboratories is severely limited by this marked method-dependency, so that a correct interpretation of analytical values is possible only for one particular analytical procedure using one type of kit.

#### Dependency of results on the date of the interlaboratory survey

Figures 4 and 5 show that kit-dependent differences were not constant during the period of the interlaboratory surveys. Thus, in the analysis of carcinoembryonic antigen during 1988 and 1989 there was a marked decrease in the differences between the medians of kit 1 76 and those of the total collective (fig.

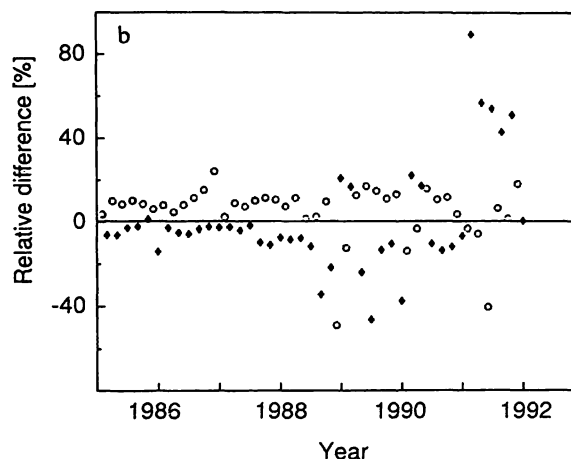
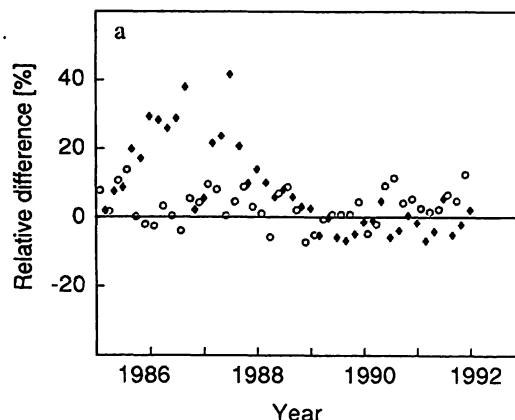


Fig. 5. a) Relative differences between the medians of the results for  $\alpha$ -fetoprotein obtained with kit 1 76 (◆) or kit 2 04 (○) and the corresponding medians of the total collective. b) Relative differences for kit 1 04 (○) and kit 1 08 (◆). Interlaboratory surveys (6 samples per year) conducted from 1986 to 1991.

4). From 1986 to 1988, the medians of the  $\alpha$ -fetoprotein determination with kit 1 76 showed relatively large differences, whereas the differences in subsequent years were very small (fig. 5a). The results from kit 1 08 remained constant with respect to the total collective until 1988, but then displayed varying differences, which since 1991 have been mostly positive compared with the total collective (fig. 5b). These phenomena suggest that changes have occurred in kit quality. Such changes would not be in the interests of patient care.

#### Repeated determinations in individual laboratories

The differences discussed so far between the interlaboratory survey results for a given analysed material are largely the differences between the medians of individual collectives of results. The predictive value of such differences may be limited for the following reasons.

- The analysed material changes between interlaboratory surveys, and resulting changes in the matrix of the samples could be a cause of inconsistencies in the observed differences.
- The composition of the participating collective is not constant, because not every laboratory takes part in every interlaboratory survey. Especially in the case of small collectives, the possibility cannot be excluded that this influences the position of the median.

Sometimes the same sample material was used in different interlaboratory surveys. In these cases, the sample matrix was not a variable, and therefore could not have been responsible for the change in position

of the medians of the collectives. Some results of such repeated determinations are shown in table 2. In addition, in order to eliminate the influence of the variable composition of the collectives, a further study was made of results obtained with the same kit in the same laboratory for the analysis of samples of the same material in different interlaboratory surveys. Examples of these results are detailed in figure 6 for the analytical quantity,  $\alpha$ -fetoprotein. Although the number of cases eligible for inclusion in this study is relatively small, the figure clearly shows that all users of kit 1 08, who participated in both interlaboratory surveys, obtained markedly higher values in June 1991 than in November 1988 for the same sample material. Moreover, the results from kit 2 04 were very non-

Tab. 2. Medians of the results for the determination of  $\alpha$ -fetoprotein (a) and carcinoembryonic antigen (b) in the same sample material in two interlaboratory surveys. Three surveys were based on previously used sample material, the corresponding first and second surveys being separated by different time intervals.

1) For all participating laboratories that used the same kit.

2) For those laboratories in 1) that participated in both surveys.

(Med = median, n = number of laboratories, D = difference between the medians of the first and second determination)

AFP* (kU/l)		Repetition interval								
Kit	Coll.	3 Years			1 Year			4 Months		
		Oct 88 Med (n)	June 91 Med (n)	D	Jan 90 Med (n)	Jan 91 Med (n)	D	June 89 Med (n)	Nov 89 Med (n)	D
1 08	1)	49.1 (17)	63.7 (14)	+14.6	12.2 (18)	19.9 (16)	+7.7	3.3 (17)	3.8 (14)	+0.5
	2)	51	64.8 (7)	+13.8	12.6	19.8 (10)	+7.2	4.4	3.5 (12)	-0.9
2 04	1)	37.3 (20)	41.9 (44)	+4.6	10.1 (22)	10.7 (43)	+0.6	6.1 (21)	6.2 (20)	+0.1
	2)	38.3	39.4 (10)	+1.1	9.9	11 (18)	+1.1	6.2	6.2 (16)	0
2 28	1)	51.4 (11)	44.1 (29)	-7.3	11.3 (20)	11 (33)	-0.3	5.5 (16)	5.9 (18)	+0.4
	2)	51.4	48.8 (7)	-2.6	11.1	11.6 (15)	+0.5	5.6	5.9 (13)	+0.3
1 76	1)	41.7 (14)	39 (24)	-2.7	10.9 (17)	9.8 (22)	-1.1	5.8 (18)	5.9 (16)	+0.1
	2)	42.7	39.9 (7)	-2.8	11.3	10.4 (8)	-0.9	5.9	5.8 (10)	-0.1

CEA** ( $\mu$ g/l)		Repetition interval								
Kit	Coll.	3 Years			1 Year			4 Months		
		Oct 88 Med (n)	June 91 Med (n)	D	Jan 90 Med (n)	Jan 91 Med (n)	D	June 89 Med (n)	Nov 89 Med (n)	D
1 04	1)	36.6 (31)	34.3 (26)	-2.3	23.8 (33)	22.9 (29)	-0.9	16.3 (36)	15.9 (26)	-0.4
	2)	36.8	35.3 (16)	-1.5	24.3	23 (23)	-1.3	16.5	16 (25)	-0.5
1 20	1)	29.4 (15)	26.9 (23)	-2.5	22.1 (13)	20.9 (26)	-1.2	15.5 (18)	15.1 (17)	-0.4
	2)		(4)		22.1	20.3 (10)	-1.8	15.5	15.2 (12)	-0.3
1 30	1)	21.4 (9)	14.8 (19)	-6.6	25.2 (17)	28 (23)	+2.8	18.4 (16)	14.6 (14)	-3.8
	2)		(3)		25.4	29 (11)	+3.6	18.3	14.1 (11)	-4.2
2 04	1)	35.6 (37)	28.3 (47)	-7.3	22.4 (32)	18.6 (50)	-3.8	16 (32)	16 (27)	0
	2)	36.7	31 (13)	-5.7	22	18.9 (22)	-3.1	15	14.5 (17)	-0.5
2 74	1)	28.8 (14)	42.3 (18)	+13.5	33.9 (17)	36.7 (18)	+2.8	18.3 (21)	18.3 (15)	0
	2)		(3)		31.9	36.8 (8)	+4.9	18.2	19 (15)	+0.8

\*  $\alpha$ -Fetoprotein

\*\* Carcinoembryonic antigen

uniform, but taken as a whole they do not suggest that the analytical properties of the kits changes between the first and second investigation. The changes in group 2 28 were less glaring than those in 1 08, but the tendency to lower values in the second determination cannot be overlooked.

To reproduce the contents of figure 6 in abridged form, the median of each first determination and the corresponding second determination were listed in table 2. With kit 1 08, the median of the second determination (65 kU/l) was 27% higher than that of the first determination (51 kU/l). Following the same procedure, data were listed from two further repeated determinations of  $\alpha$ -fetoprotein (tab. 2), and from three repeated determinations of carcinoembryonic antigen (tab. 2b). As mentioned above, tables 2a and 2b also contain the medians for the first and second determinations of the complete subcollectives. Since they originated at least partly from different laboratories, many results in these complete subcollectives were not true repeated determinations. In most cases, however, their medians largely corresponded to those of the smaller groups of selected laboratories. At least when the medians of the first and second determinations differed by 20% and more, such differences can hardly be explained by random effects or peculiarities of the sample matrix. It is possible that, unknown to the user, the analytical method had been changed. Such changes would probably affect the reliability of the test as a diagnostic tool.

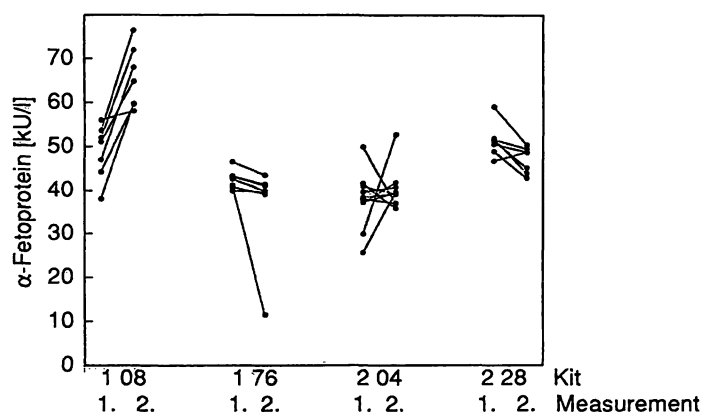


Fig. 6. Individual values from laboratories that used the same kit in two interlaboratory surveys (Oct. 1988 and June 1991), in which the same control material was analysed. The results of the first and second determination (●) are joined by a line (●—●).

The precision of the determinations is particularly crucial in follow-up studies of tumour patients, in which laboratory tests are separated by relatively long periods. Also, the "triple diagnostic", developed in recent years for the calculation of *Down* syndrome risk factors (5), uses the quantities oestradiol,  $\alpha$ -fetoprotein and human chorionic gonadotropin, and bases its predictive value (which is not undisputed) on a very good repeatability of results (6). It is therefore important that kits and the quality of their performance remain unchanged.

## Conclusion

The examples presented here of interlaboratory surveys of tumour markers show once again that analytical results can differ to a greater or lesser extent from one another, depending on the reagent combinations used for the analyses. In view of the complexity of the analytical procedures, this is a plausible state of affairs, and it has been repeatedly reported and expounded (2, 7–11). Nevertheless, there are grounds for believing that analysts and physicians are not yet sufficiently aware of its implications. This awareness is important, because the situation cannot be expected to improve in the foreseeable future.

The situation will also not be changed by the fact that international standard preparations are available for some tumour markers. The heterogeneity of the antigens, the multiplicity of antibodies and differences in the methodology of kits will probably continue to stand in the way of the desired harmonization of analytical results from different laboratories. For the time being, it will be necessary to accept a variable degree of kit dependency in the results for tumour markers. In the interests of diagnosis, it is therefore all the more important to guarantee the reproducibility of the results at least within each laboratory. However, as shown by the data for repeated determinations, this requirement was not always satisfied. For the continuous monitoring of the constancy of an analytical system, however, the conduct of interlaboratory surveys at relatively large time intervals is of only limited value. Rapid and reliable information on the repeatability of an analytical method can only be obtained by strict internal laboratory control of precision.

## References

1. Wagener, C. & Breuer, H. (1980) Diagnostische Bedeutung von Tumormarkern in der Klinischen Chemie: Bericht über die Kleinkonferenz der Deutschen Gesellschaft für Klinische Chemie vom 15.–17. November 1979 in Schloß Auel. *J. Clin. Chem. Clin. Biochem.* 18, 821–827.
2. Röhle, G., Voigt, U., Kruse, R., Oberdörster, W. & Breuer, H. (1982) External Quality Control of Determinations of Tumor-Associated Antigens. In: "Radioimmunoassay and Related Procedures in Medicine 1982" Herausgeber: International Atomic Energy Agency, Wien, 645–652.

3. Röhle, G. & Voigt, U. (1986) Concept for the Evaluation of Analytical Results in Clinical Chemistry. Part 1: Concentration Dependent Dispersion Profiles of Survey Results. *Commun. Lab. Med.* 2, 81–85.
4. Neumaier, M., Fenger, U. & Wagener, C. (1985) Monoclonal Antibodies for Carcinoembryonic Antigen (CEA) as a Model System: Identification of Two Novel CEA-related Antigens in Meconium and Colorectal Carcinoma Tissue by Western Blots and Differential Immunoaffinity Chromatography. *J. Immunol.* 135, 3604.
5. Wald, N. J., Cuckle, H. S., Densem, J. W., Nanchahal, K., Royston, P., Chard, T., Haddow, E., Knight, G. J., Palomaki, G. E. & Canik, J. A. (1988) Maternal Serum Screening for Down's syndrome in Early Pregnancy. *Br. Med. J.* 297, 883–887.
6. Reynolds, T. M. (1992) Practical Problems in Down Syndrome Screening. *Commun. Lab. Med. (continued series)* 1, 31–38.
7. Fateh-Moghadam, A., Hero, B., Stieber, P., Nagel, D. & Jaworek, D. (1990) Investigation of Different CA 19-9 Assays: Clinical Significance and Consequences for the Interpretation, the Assessment of Multicenter Trial Results and the Standardisation (Meeting Abstract). *J. Tumor Marker Oncol.* 5, 224.
8. Dahlmann, N. (1991) Qualitätssicherung bei der Bestimmung von Tumormarkern: Darstellung und Diskussion eines Ringversuchs. *Lab. Med.* 15, 344–346.
9. Oosterom, R. & Stoter, G. (1992) Standardizing Human Chorionic Gonadotropin Measurements in Serum of Testicular Cancer Patients. *Clin. Chem.* 38, 601.
10. Bauer, H. W. (1992) Stellenwert des Prostata-spezifischen Antigens (PSA) für Therapie und Verlaufskontrolle des Prostatakarzinoms. *Urologe (Ausg. B)* 32, 24–28.
11. Lamerz, R. (1992) Allgemeine Kriterien zur Bestimmung von Tumormarkern im Labor und Anwendung in der Klinik. *Lab. Med.* 16, 13–20.

Prof. Dr. G. Röhle  
Institut für Klinische Biochemie  
Sigmund-Freud-Straße 25  
W-5300 Bonn 1  
Germany